



The Children Depression Inventory (CDI) and the Beck Depression Inventory (BDI): Their validity as screening measures for major depression in a group of Puerto Rican adolescents¹

Carmen L. Rivera², Guillermo Bernal, and Jeannette Rosselló
(University of Puerto Rico, Puerto Rico)

(Recibido 15 de octubre 2004/ Received October 15, 2004)

(Aceptado 14 de enero 2005 / Accepted January 14, 2005)

ABSTRACT. This quasi-experimental study evaluates the efficiency of the Children Depression Inventory (CDI) and the Beck Depression Inventory (BDI) in their Spanish versions, as screening measures for Major Depressive Disorder (MDD) with a Puerto Rican clinical sample of adolescents. The sample consisted of 130 adolescents between 13 to 18 years of age. The results obtained show that the best cut-off point for the CDI to identify MDD is 20 with a sensibility index of .69, specificity of .43, a positive predictive value of .64, and negative predictive value of .49. The best cut-off point for the BDI is 12 with a sensibility of .65, specificity of .50, a positive predictive value of .67, and a negative predictive value of .47. In conclusion, the CDI and the BDI are fairly good instruments to be used in the screening for MDD for clinical samples with Puerto Rican adolescents. Specifically for the CDI, the results obtained differ considerably from the cut-off points, sensitivity and specificity scores obtained by Kovacs. Even when using the cut-off points recommended by the author of the instrument, the sensitivity and specificity scores to identify a disorder differ from one population to another; therefore, we cannot assume metric equivalency.

¹ Work on this paper was in part supported by a grant from the National Institute of Mental health R24-MH49368. The authors are grateful to Dr. Rafael Ramírez for his statistical advice, to Amy Leah Fontenot for her editorial recommendations and to the research and administrative staff of the University Center for Psychological Services and Research at the University of Puerto Rico, Río Piedras Campus.

² Correspondence: CUSEP–University Center for Psychological Services and Research. PO Box 23174. San Juan 00931-3174 (Puerto Rico). E-mail: crivera4@rrpac.upr.clu.edu

KEY WORDS. Major Depressive Disorder. Adolescents. CDI. BDI. Quasi-experimental study.

RESUMEN. Este estudio cuasi-experimental evalúa la eficiencia del Inventario de Depresión en Niños (CDI) y el Inventario de Depresión de Beck (BDI), en sus versiones en español, como medidas de cernimiento para el trastorno de depresión mayor (TDM) en una muestra clínica de adolescentes puertorriqueños. La muestra incluía 130 adolescentes con edades entre 13 y 18 años. Los resultados obtenidos demostraron que el mejor punto de corte para identificar TDM fue una puntuación de 20 con un índice de sensibilidad de 0,69, especificidad de 0,43, un valor predictivo positivo de 0,64 y un valor predictivo negativo de 0,49. El mejor punto de corte para el BDI fue una puntuación de 12 con un índice de sensibilidad de 0,65, especificidad de 0,50, un valor predictivo positivo de 0,67 y un valor predictivo negativo de 0,47. Se puede concluir que el CDI y el BDI pueden ser útiles como instrumentos de cernimiento para identificar TDM en muestras clínicas de adolescentes puertorriqueños. Los resultados obtenidos difieren de los puntos de corte e índices de sensibilidad y especificidad obtenidos por Kovacs para el CDI. Los resultados evidencian que la sensibilidad y especificidad para identificar un trastorno pueden diferir de una población a otra por lo cual no se puede asumir equivalencia psicométrica.

PALABRAS CLAVES. Trastorno depresivo mayor. Adolescentes. CDI. BDI. Estudio cuasi-experimental.

RESUMO. Este estudo quase-experimental avalia a eficiência do Inventário de Depressão em crianças (CDI) e o Inventário de Depressão de Beck (BDI), nas suas versões em espanhol, como medidas discriminativas para a perturbação de depressão maior (TDM) numa amostra clínica de adolescentes porto-riquenhos. A amostra incluía 130 adolescentes com idades entre 13 e 18 anos. Os resultados obtidos demonstram que o melhor ponto de corte para identificar TDM foi uma pontuação de 20 com um índice de sensibilidade de 0,69, especificidade de 0,43, um valor preditivo positivo de 0,64 e um valor preditivo negativo de 0,49. O melhor ponto de corte para o BDI foi uma pontuação de 12 com um índice de sensibilidade de 0,65, especificidade de 0,50, um valor preditivo positivo de 0,67 e um valor preditivo negativo de 0,47. Pode-se concluir que o CDI e o BDI podem ser úteis como instrumentos de triagem para identificar TDM em amostras clínicas de adolescentes porto-riquenhos. Os resultados obtidos diferem dos pontos de corte e índices de sensibilidade e especificidade obtidos por Kovacs para o CDI. Os resultados evidenciam que a sensibilidade e a especificidade para identificar uma perturbação podem diferir de uma população para outra, pelo que, não se pode assumir equivalência psicométrica.

PALAVRAS CHAVE. Perturbação depressiva maior. Adolescentes. CDI. BDI. Estudo quase-experimental.

Introduction

The development and validation of measures to evaluate depression in adolescents has lagged behind the theoretical and clinical developments for these populations. The absence of information on the validity and reliability of instruments is particularly acute in Latino populations (Diez Zamorano, 2003). The use of screening measures developed in the United States or in other countries is widely accepted among Latino/a professionals in general and Puerto Rican mental health professionals in particular. The validation of instruments across different cultural groups is crucial to provide evidence of the scale sensibility to cultural diversity, and to help identify symptomatic differences between groups (Abdel-Khalek and Soliman, 1999). This draws the attention to the situation of metric equivalence defined by Okasaki and Sue (1995) as the assumption that a given metric has the same meaning in different cultures, and reported research showing the dangers of assuming metric equivalence of translated measures. Added to the problem of a lack of sensibility to cultural diversity, there is also a definite lack of strong evidence of the capabilities of the depression measures scales to screen for Mood Disorders, specifically Major Depressive Disorder (MDD), in adolescents. According to Roberts, Lewinsohn and Seeley (1991), even though non-clinical measures, usually in the form of brief symptoms check lists, are widely used for case ascertainment in psychiatric research, there is little evidence on the concordance between screening scales and diagnostics procedures. The information provided by the screening scales is critical for the decision making process, as it enables practitioners to differentiate possible cases from non-cases, and to identify those cases that may need immediate intervention.

The Children's Depression Inventory (CDI) (Kovacs, 1992) is consistently cited in the literature as a good instrument to identify depressive symptoms in children and adolescents (Chan, 1997; Fristad, Emery, and Beck, 1997; Nurcombe, Seifer, Scioli, Tramontana, Lexington, and Beauchesne, 1989). Over three fourths of the studies utilizing self-report measures utilized the CDI. However, although the CDI is a good indicator of self-reported distress in children and has being reported as a good screening measure for depression (Kovacs, 1983; Nurcombe *et al.*, 1989), some studies have demonstrated that it does not have the adequate sensitivity and specificity as a screening measure of depression (Fristad, Weller, Weller, Teare, and Preskorn, 1991; Nelson, Politano, Finch, Wendel, and Mayhall, 1987; Saylor, Finch, Spirito, and Bennett, 1984; Weiss, Weisz, Politano, Carey, Nelson, and Finch, 1991). A key issue in assessing the predictive validity of an instrument for a specific disorder is in terms of where to establish the cut-off point. For the CDI, Kovacs (1992) suggests the use of a cut-off point of 12 or 13 for homogeneous samples (e.g., clinical). A lower cut-off point is supported by several studies cited by Kovacs (Garvin, Leber, and Kalter, 1991; Kazdin, Colbus, and Rodgers, 1986; Lobovits and Handal, 1985) where higher rates of depression would be expected. However, in a study with 621 Chinese adolescents, Chan (1997) found that a cut-off point of 20 was a better option to identify depressive symptoms in his sample. Another option is presented by Donnelly (1995) and Hessel and Matson (1984) who selected depressed subjects on the basis of one standard deviation greater than the mean CDI. Mausner and Kramer (1985) are of the opinion that the decision of establishing a cut-

off point should be determined by: the cost of diagnostic testing of false positives; the importance of not missing a possible case; the likelihood that the population will be re-screened at a reasonable interval; and the prevalence of the disease. Those authors believe that adequate indexes can be obtained only when the prevalence rate is above 15 or 20 percent. Therefore, those indexes will depend a great deal upon the population under study. This draws the attention to the psychometric issue raised by Rogler (2000) concerning the use of scales of whether the scores achieved by members of one cultural group have the same meaning as those achieved by members of another cultural group.

On the other hand, although initially designed to assess the severity of depressive symptoms in the adult population, the BDI is another scale used to assess depressive symptoms in adolescents (Ambrosini, Mets, Bianchi, Rabinovich, and Undie, 1991; Barrera and Garrison-Jones, 1988; Beck, Rush, Shaw, and Emery, 1979). The BDI has become widely accepted in clinical psychology and psychiatry for assessing the intensity of depression in psychiatric patients (Piotrowski, Sherry, and Keller, 1985). However, Ambrosini *et al.* (1991) consider the BDI as one of the most widely used self-administered depression inventories whose psychometric properties have been well investigated, but only for adult populations.

Beck and Steer (1993) recognize that there is considerable debate over the use of the BDI for screening purposes. Beck and Steer cited the works of Gotlib (1984) and Tanaka-Matsumi and Kameoka (1986), which state that high BDI total scores for university students were not necessarily indicative of depression, but represented overall adjustment problems. In a study carried out by Strober, Green, and Carlson (1981) with psychiatrically hospitalized adolescents, a cut-off point of 16 on the BDI correctly identified 81% of the sample with MDD, with a false positive rate of 14% and a false negative rate of 5%. However, several studies have demonstrated the inconsistencies of the BDI as a screening measure with adolescents (Ambrosini *et al.*, 1991; Barrera and Garrison-Jones, 1988; Kashani, Sherman, Parker and Reid, 1990; Strober *et al.*, 1981). For example, the psychometric results generated in a study by Ambrosini *et al.* (1991) validate the utility of the BDI for distinguishing adolescents who are syndromally depressed from non-depressed adolescents in an outpatient setting. Meanwhile, Roberts *et al.*, (1991) propose that although there appears to be evidence for the content, concurrent and construct validity of the BDI, the question of how well the scale predicts clinical depression in adolescents still remains.

As it may be apparent, in all of the studies cited above for the CDI or the BDI Latino samples were not the main sample represented. Furthermore, a mere translation and adaptation of the scale is not enough to consider the scale adequate for the Latino population, since the psychometrics properties extend beyond the issue of language, which is just one characteristic of the sample. An epidemiological study in Puerto Rico revealed a 5.9% prevalence rate of MDD in children and adolescents between 4 to 16 years old (Bird *et al.*, 1988). The authors found that depression and dysthymia were the third most frequent disorder in Puerto Rican children and adolescents. In a study with 146 children and adolescents from three outpatient mental health clinics and a residential drug treatment facility in Puerto Rico, to test-retest the reliability of the Spanish Diagnostic Interview Schedule for Children (DISC-IV), 21.9% received a classification of depressive

diagnosis (Bravo *et al.*, 2001). When considering those between the ages of 11 to 17 years of age, 24.7% received a diagnosis of depression. In Puerto Rico, both, the CDI and the BDI have been previously studied, though not with the intention to assess the cut-off point of the scales. The CDI was translated and culturally adapted to a Spanish version considering the semantic, contextual and technical equivalency to the English version (Rosselló, Guisasola, Ralat, Martínez and Nieves, 1992). Bernal, Rosselló, and Martínez (1996) evaluated the psychometric properties of the Spanish version of the CDI with both a Puerto Rican community and a clinical sample of adolescents whose ages ranged from 13 to 20 years. For community samples (265 participants) the internal consistency of the CDI was .82 and .80 for the clinical sample (59 participants). Bonilla, Bernal and Varas (1997) translated the BDI, adapting it to a Spanish version and adding one item to the original English version in order to comply with the diagnostic criteria for major depression according to the DSM-IV. This modified version was administered to 351 students and reported an internal consistency of .88 for the instrument.

However, for the CDI as well as for the BDI in their Spanish versions, the authors kept the cut-off points suggested by Beck and Steer (1993) and Kovacs (1992) in the original English versions. To our knowledge, neither the CDI nor the BDI have been studied in terms of their validity as a screening measure for depressive diagnosis with Puerto Rican adolescents. The objective of this quasi-experimental study (Montero and León, 2005) is to evaluate the efficiency of two commonly used instruments to measure depressive symptoms with a Puerto Rican clinical sample. The main question of this research is how well the Spanish translations and adaptations of the CDI and the BDI work as screening measures for MDD in an adolescent clinical sample, as measured by their predictive validity indexes. The current paper followed the guidelines proposed by Ramos-Alvarez and Catena (2004).

Method

Participants

The sample consisted of 130 Puerto Rican (Spanish speaking) adolescents originally screened to be part of a clinical trial to assess the effectiveness of two treatments on depression in adolescents. Participants eligible for this study were between 12 and 18 years of age and met regular DSM-IV criteria (American Psychiatric Association, 1994) for Major Depression Disorder (MDD) and/or Dysthymic Disorder (DD) and were deemed by a clinical interviewer to be impaired. Also eligible were those who did not meet the DSM-IV criteria for MDD but obtained a score of 13 or higher in the CDI as suggested by Kovacs (1992) for moderate to severe depressive symptoms. Since the sample was part of a clinical trial, the following conditions resulted in exclusion from the study: serious imminent suicide risk, psychosis, mental retardation, hyper-aggression, current regimen of psychopharmacological medication or psychotherapy, involvement in legal proceedings, bipolar disorders, conduct disorder, or drug use. Once a written consent form was obtained from their parents, and cosigned by the adolescents, the adolescents were interviewed and evaluated for MDD. After the interview, the adolescents completed the CDI and BDI inventories as part of a battery of instruments to evaluate

depressive symptoms and other related conditions. Adolescents were referred from schools in the Metropolitan San Juan area. Three hundred and twenty-two referrals were received; 130 met criteria for inclusion in the study. The participants were 54% (70) females and 46% (60) males between the ages of 13 and 18, with a mean age of 15. Their mean grade level was 9.28, with 50.8% of the participants attending public schools and 49.2% attending private schools.

Instruments

- The Children’s Depression Inventory (CDI) is a 27 item self-rated symptom oriented scale translated to a Spanish version and validated by Rosselló *et al.* (1992). In its original English version, this instrument was developed based on the Beck Depression Inventory (BDI), to distinguish youth with the psychiatric diagnoses of major depressive or dysthymic disorder from those with other psychiatric conditions or non-selected “normal” schoolchildren (Kovacs, 1992). Each item is scored according to one of the following alternatives: 0 = absence of the symptom, 1 = moderate symptom, 2 = severe symptom. For the present study, a reliability coefficient of .82 was obtained for the CDI.
- The Beck Depression Inventory (BDI) is a 22-item Spanish version that was modified from the original English version (21 items) to consider the diagnostic criteria for major depression according to the DSM-IV (Bonilla, Bernal, Santos, and Santos, 2004). These investigators reported high internal consistency. This is a well-known and widely used instrument for the evaluation of depressive symptoms (Beck and Steer (1984); Beck, Ward, Mendelson, Mock, and Erbaugh, 1961). Symptoms include vegetative, mood, and cognitive aspects of depression. The items are score from 0 to 3: 0 = symptom not present, 1 = symptom is present, 2 = symptom is moderate, 3 = symptom is severe. For this study, the BDI obtained a reliability coefficient of .91. According to their manuals for the original English versions, the cut-off points to identify moderate or severe depressive symptoms should be 17 for the BDI and 12 for the CDI. However, Kovacs (1992) as well as Beck and Steer (1993) recognized that the cut-off points should be set depending on the false negatives and false positives that the assessor wants to allow in his data.
- Diagnostic Interview Schedule for Children (DISC-2.3). The NIMH DISC version 2.3 is a highly structured research diagnostic instrument that assesses the most common DSM-III-R diagnoses among children and adolescents. A translated version of this instrument was used to provide a structured format for the interview and reduce informant variance in the interviews with the adolescents and their parents (Bravo, Canino, Rubio-Stipec, and Woodbury-Fariña, 1991). However, the diagnoses of MDD were not computer derived from the DISC algorithms. Only the section that evaluates affective disorders was used.
- A Symptoms Checklist based on DSM-IV criteria for MDD was completed for each participant.

Procedure for the diagnosis

Doctoral candidates in clinical psychology performed the clinical evaluations. A Ph.D. clinical psychologist supervised all evaluations. To assess clinical depression at intake, the depression section of the Spanish version of Diagnostic Interview Schedule for Children was used (adapted to a Spanish version by Bravo *et al.*, 1991). The interviewers were aware of the principal aim of the clinical trial, but not of the aims of the present study. DSM-IV diagnoses for MDD were made by a clinician using the information obtained by the DISC and the symptoms checklist based on DSM-IV obtained for each participant. According to the DSM-IV criteria, 62.5% of the adolescents received a diagnosis of MDD, of whom 50.7% also met the criteria for dysthymia, 5.0% received a diagnosis of dysthymia, and 32.5% did not comply with the criteria for either MDD or dysthymic disorder. Those rates are reasonable, considering that the participants were referred to an outpatient clinic for depressive symptomatology.

Analysis of data

Once the diagnosis of MDD was established, the following information was calculated:

- True positives (TP). Those who have MDD and score in the range of severe or moderate depression on the instruments (equal to, or above, the cut-off point).
- True negatives (TN). Those who score negative for MDD and score below the cut-off points on the instruments.
- False positives (FP). Those who score above the cut-off points on the instruments but do not comply with the DSM-IV criteria for MDD.
- False negatives (FN). Those who score below the cut-off points on the instruments but have MDD.

With the above information the following scores were calculated:

- Sensitivity. Capacity of the instruments to correctly identify those true positives expressed as the percentage of cases scoring positive from those diagnosed with MDD ($SEN = TP/(TP + FN)$).
- Specificity. Capacity of the instruments to correctly identify those true negatives expressed as the percent of cases scoring negatives from those non diagnosed with MDD ($SPE = TN/(TN + FP)$).
- Positive Predicted Power. The probability of having MDD given the presence of a positive result expressed as the percent of cases with MDD from those who score positive on the scale ($PPP = TP/(TP + FP)$).
- Negative Predicted Power. The probability of not having MDD given the absence of a positive result expressed as the percent of true negative from those who score negative on the scale ($NPP = TN/(TN + FN)$).
- Total Predictive Power. Overall capability of the instrument to accurately identified either cases or non-cases for a given cut-off point ($TPP = TP + TN/(TP + FP + FN + TN)$).

The test efficiency indexes for the CDI and the BDI were calculated at different cut-off points for each instrument including those suggested by the manuals in their original English versions. The cut-off points were evaluated considering the SEN, SPE, PPP, NPP, and TPP indexes as well as the TP, TN, FP, and FN. For the purposes of this paper, obtaining a balance between the indexes as well as between the FP and FN rates was important when establishing the cut-off points for the scales. The correlation between the instruments was moderately high (.69 at .01 significance level), suggesting that both instruments are measuring the same construct.

Results

Evaluating the psychometric characteristics of the CDI at various cut-off points, the results obtained show that the best cut-off point for identifying MDD in a clinical sample of adolescents may differ depending on the purposes of the clinician or the researcher. For example, for mere screening purposes, a cut-off point of 15 could be better, keeping a lower rate of FN (see Table 1) with a SEN index of .88, SPE of .13, a PPP of .60, NPP of .44 and TPP of .58. However, for clinical trials (which is our case) or research purposes, for which the rate of FP and FN should be kept low and a balance should be kept between them, the cut-off point of 20 and above would be recommended, with a prevalence rate of 64, a SEN index of .69, SPE of .43, a PPP of .64, NPP of .49 and TPP of .58. Above this point, even though the FP rates continue to drop, the TP rates also decrease substantially and the FN rates increase. As can be seen, above a cut-off point of 20 the TPP index of the scale also decreases.

TABLE 1. Predictive values, diagnostic efficiency and percentage distribution of the MDD at different cutoff points of the CDI.

<i>Cutoff points</i>	<i>Q %</i>	<i>SEN</i>	<i>SPE</i>	<i>PPP</i>	<i>NPP</i>	<i>TPP</i>	<i>TP %</i>	<i>FP %</i>	<i>FN %</i>	<i>TN %</i>
CDI										
12	95	.97	.09	.61	.71	.62	57.7	36.9	1.5	3.8
13	92	.92	.09	.60	.46	.58	54.6	36.9	4.6	3.8
14	90	.91	.11	.60	.46	.58	53.8	36.2	5.4	4.6
15	88	.88	.13	.60	.44	.58	52.3	35.4	6.9	5.4
16	82	.82	.17	.59	.39	.55	48.5	33.8	10.8	6.9
17	78	.79	.25	.60	.45	.57	46.9	30.8	12.3	10.0
18	74	.78	.32	.63	.50	.59	46.2	27.7	13.1	13.1
19	70	.74	.36	.63	.49	.58	43.8	26.2	15.4	14.6
20	64	.69	.43	.64	.49	.58	40.8	23.1	18.5	17.7
21	57	.60	.47	.62	.45	.55	35.4	21.5	23.8	19.5
22	52	.55	.51	.62	.44	.53	32.3	20.0	26.9	20.8
23	45	.47	.57	.61	.42	.51	27.7	17.7	31.5	23.1
24	38	.42	.68	.65	.44	.52	24.6	13.1	34.6	27.7

Q= Score Prevalence *SEN*= Sensitivity; *SPE*= Specificity; *PPP*= Positive Predictive Power; *NPP* = Negative Predictive Power; *TPP* = Total Predictive Power; *TP* = True Positive; *FP* = False Positive; *FN* = False Negative; *TN* = True Negative

For the BDI scale, the results obtained to identify MDD for clinical trials show quite a low cut-off point of 12 or above (see Table 2), with a prevalence rate of 59, a SEN index of .65, SPE of .50, a PPP of .67, NPP of .47 and TPP of .59. Since the FN rates for the BDI appear to be higher than for the CDI, it would be expected that the cut-off point recommended for screening purposes would be substantially lower than 12, obtaining highly questionable results. Kappa coefficients were obtained for this analysis as well as for the CDI. However, since the results obtained were quite erratic, and given that this test has triggered controversies concerning its credibility or utility as an agreement measure, these results were not included. For example, in this study, some of those cut-off points with the highest kappa coefficients were the ones with the highest FN rates or cut-off points with considerably low kappa coefficients obtained the best psychometrics properties.

TABLE 2. Predictive values, diagnostic efficiency and percentage distribution of the MDD at different cutoff points of the BDI.

<i>Cutoff points</i>	<i>Q %</i>	<i>SEN</i>	<i>SPE</i>	<i>PPP</i>	<i>NPP</i>	<i>TPP</i>	<i>TP %</i>	<i>FP %</i>	<i>FN %</i>	<i>TN %</i>
BDI										
9	74	.75	.27	.62	.41	.56	46.0	28.2	15.3	10.5
10	71	.75	.35	.65	.47	.60	46.0	25.0	15.3	13.7
11	65	.71	.44	.67	.49	.60	43.5	21.8	17.7	16.9
12	59	.65	.50	.67	.47	.59	39.5	19.4	21.8	19.4
13	57	.61	.50	.67	.44	.56	37.1	19.4	24.2	19.4
14	54	.58	.52	.66	.44	.56	35.5	18.5	25.8	20.2
15	50	.53	.54	.65	.42	.53	32.3	17.7	29.0	21.0
16	47	.50	.58	.66	.42	.53	30.6	16.1	30.6	22.6
17	43	.49	.67	.70	.45	.56	29.8	12.9	31.5	25.8
18	38	.43	.71	.70	.44	.54	26.6	11.3	34.7	27.4
19	36	.42	.75	.73	.45	.55	25.8	9.7	35.5	29.0
20	34	.39	.75	.71	.44	.53	24.2	9.7	37.1	29.0
21	31	.38	.81	.76	.45	.55	23.4	7.3	37.9	31.5

Q= Score Prevalence *SEN*= Sensitivity; *SPE*= Specificity; *PPP*= Positive Predictive Power; *NPP* = Negative Predictive Power; *TPP* = Total Predictive Power; *TP* = True Positive; *FP* = False Positive; *FN* = False Negative; *TN* = True Negative

Discussion

The CDI and the BDI are fairly good instruments to be used in the screening for MDD for clinical samples with Puerto Rican adolescents. For the CDI, a cut-off point of 20 to identify MDD is much higher than the score of 13, as suggested by Kovacs (1992). Actually, selecting the cut-off point suggested by Kovacs seems to be more appropriate for mere general screenings where FN rates are expected to be lower and FP rates higher. However, this score is compatible with the score of 19 suggested by Kovacs to identify severe depression in more heterogeneous samples, such as a community or school sample. The results corroborate the Donnelly (1995), Chan (1997) and Helsen and Matson (1984) studies, which recommend higher cut-off points for the CDI to more

accurately identify depressive symptoms in adolescent samples. In addition, the results obtained differ considerably from the sensitivity and specificity scores obtained by Kovacs (1983) for the cut-off points of 12 or 13. Even though the sensitivity scores reported here were higher than the moderate scores obtained by Kovacs, our specificity scores at those cut-off points were much lower. These results suggest that, even when using the same cut-off points recommended by the author of the instrument, the sensitivity and specificity scores may differ from one population to another.

The rate of FP (23%) using the 20 cut-off point could be considered high. However, a probability of 64% (PPP) to correctly identifying MDD, given the presence of a positive result on the scale on a screening process, may justify such a rate. After all, the objective of a screening instrument is to identify probable diagnosis, not as a substitute of a full diagnostic procedure. Furthermore, receiving a preliminary diagnosis of MDD does not have to be negative for the participant. The literature has demonstrated the importance of paying attention to the FP in depression studies, since those cases could be reflecting a depressive syndrome, dysthymia, or another psychopathology or social dysfunction (Compas, Ey, and Grant, 1993; Gotlib, Lewinsohn, and Seeley, 1995; Hammen and Rudolph, 1996; Joiner Jr., Schmidt, and Schmidt, 1996; Nurcombe *et al.*, 1989).

For the BDI, the cut-off point of 12 obtained is lower than the score of 17 suggested by other researchers (Barrera and Garrison-Jones, 1988; Beck and Steer, 1993; Kashani *et al.*, 1990; Strober *et al.*, 1981). The results also differ from Marton, Churchard, Kutcher, and Koremblum's study (1991), which reported that the patients with an affective disorder had higher scores on the BDI than did patients with other disorders. They also differ from the results obtained by Roberts *et al.* (1991) in a community sample in which the PPP and NPP for the BDI were .10 and .99 respectively, practically disqualifying the instrument as a screening measure. However, Marton and colleagues suggest that the instrument can be expected to vary in its utility depending on the demographic and diagnostic characteristics of the patient population. Furthermore, they believe their study provides evidence that the utility of the BDI varies with the adolescent population being studied and that it may be less useful in more typical populations with lower base rates of depression.

Barrera and Garrison-Jones (1988) believe that the utility of the BDI (and we may add, any other screening measure) will depend upon some validity issues. In clinical settings, depression-screening measures are valuable if they result in high sensitivity. For use in clinics, low specificity is less critical because false positives presumably would be detected when high scores on the screening measure are referred for more extensive assessment. In research settings, there are different considerations because the screening measures are needed to identify research subjects who are comparable to individuals who would receive psychiatric diagnoses. Therefore, it is critical to guard against false positives so that the identified "depressed" groups do not include subjects who wouldn't meet diagnostic criteria for depression.

Nevertheless, the issue of TP, TN, FP, and FN rates is compounded in a research setting that also provides clinical services. In this case, capacity to identify the TP and TN is critical for a more efficient use of the resources available and the implications

it has for the project, as well as for the treatment or cost. An adequate identification of the probable cases and non-cases, at the screening stage, maximizes the efficiency of the resources more efficiently. The capacity to properly identify cases with a screening instrument is also cost-effective for the assessment and documentation of the efficiency of either psychotherapeutic or pharmacological treatments. However, it is equally important to decrease the rate of FN, as these cases should not be discarded not only from the study, but from the possibility of receiving the treatment they need.

In clinical terms, the proper identification of MDD in adolescents versus the non-depressed or those who have dysthymia allows for the design of a more efficient therapeutic model. Without eliminating the clinical interview, essential in every assessment, an adequate screening instrument provides basic information from the beginning of the assessment stage.

It should be noted that this sample source is a limitation for the generalizability of the results to non-clinical Puerto Rican adolescents. Another limitation is that this study does not include the clinical approach to assessment of other mental health disorders that are known to co-occur with depression. However, the results with the CDI and the BDI suggest that both instruments can be used to identify symptomatology related to MDD in Puerto Rican adolescents in clinical settings. There is still the need for a similar study in a community sample of adolescents. However, the sample needed for such a study must be substantially higher in order to obtain a high enough prevalence of MDD cases to evaluate the psychometric properties of the scale. These instruments may be useful in identifying symptoms of depression in adolescents who may be in need of assessment and treatment in other non-clinical settings. As Helsen and Matson (1984) contend, the use of different cut-off points has its limitation for comparison of results, establishment of prevalence of the condition, or validity of the instrument with other samples. We believe that it is better to think in terms of how well the instruments fit our sample, rather than how well the sample fits the instrument for comparative purposes with other samples. Perhaps this strategy may limit our capacity to compare our results to other samples. Yet, this strategy insures that the instruments will be used responsibly with our adolescent population.

References

- Abdel-Khalec A.M. and Soliman H.H. (1999). A cross cultural evaluation of depression in children in Egypt, Kuwait, and the United States. *Psychological Reports*, 85, 973-980.
- Ambrosini, P.J., Mets, C., Bianchi, M., Rabinovich, H., and Undie, A. (1991). Concurrent validity and psychometric properties of the Beck Depression Inventory in outpatient adolescents. *Journal of American Academy of Child Adolescents Psychiatry*, 30, 51-57.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of mental Disorders (DSM-IV)* (Fourth Edition). Washington, DC: American Psychiatric Association.
- Barrera, M. Jr. and Garrison-Jones, C.V. (1988). Properties of the Beck Depression Inventory as a screening instrument for adolescent depression. *Journal of Abnormal Child Psychology*, 16, 263-273.
- Beck, A. T., Rush, A.J., Shaw, B. F., and Emery, G. (1979). *Cognitive therapy of depression*. New York: Guildford Press.

- Beck, A.T. and Steer, R.A. (1984). Internal consistencies of the original and revised Beck Depression Inventory. *Journal of Clinical Psychology*, 40, 1365-1367.
- Beck, A.T. and Steer, R.A. (1993). *Beck Depression Inventory: Manual*. San Antonio: The Psychological Corporation, Hartcourt Brace & Company.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Bernal, G., Rosselló, J., and Martínez, A. (1996). El inventario de depresión para niños y niñas: propiedades psicométricas en dos muestras Puertorriqueñas. *Revista de Psicología Contemporánea*, 4, 12-23.
- Bernal, G., Varas, N., Bonilla, J. Santos, A., and Maldonado, M. (1997, november). *Depresión y género: Resultados iniciales de un estudio colaborativo en el Caribe sobre los estados de ánimo*. Presented at the Congress 30 años de Psicología Dominicana: Pasado, presente y futuro. Santo Domingo.
- Bird, H. R., Canino, G., Rubio-Stipec, M., Ribera, J., Sessman, M., Woodburry, M., Huertas-Goldman, S., Pagán A., Sánchez, L. A., and Moscoso, M. (1988). Estimate of the prevalence of childhood maladjustment's in a community survey in Puerto Rico. *Archives of General Psychiatry*, 45, 1120-1126.
- Bonilla, J., Bernal, G., Santos, A., and Santos, D. (2004). A revised Spanish version of the Beck Depression Inventory: Psychometric Properties with a Puerto Rican Sample of College Students. *Journal of Clinical Psychology*, 60, 1, 119-130.
- Bonilla, J. Bernal, G., and Varas, N. (1997). *Propiedades psicométricas de una versión modificada del Inventario de depresión de Beck que considera los criterios para la depresión mayor de acuerdo al DSM-IV*. Unpublished Manuscript.
- Bravo, M., Canino, G.J., Rubio-Stipec, M., and Woodbury-Fariña, M., (1991). A cross-cultural adaptation of a psychiatric epidemiologic instrument: The Diagnostic Interview Scheduled's Adaptation in Puerto Rico. *Culture, Medicine and Psychiatry*, 15, 1-18.
- Bravo, M., Ribera, J., Rubio-Stipec, M., Canino, G.J., Shrout, P., Ramírez, R., Fabregas, L., Chavez, L., Alegría, M., Bauermeister, J., and Martínez, A. (2001). Test-retest reliability of the Spanish Version of the Diagnostic Interview Scheduled for Children (DISC-IV). *Journal of Abnormal Child Psychology*, 29, 433-444.
- Chan, D.W., (1997). Depressive symptoms and perceived competence among Chinese secondary school students in Hong Kong. *Journal of Youth and Adolescence*, 26, 303-319.
- Compas, B. E., Ey, S., and Grant, K.E. (1993). Taxonomy, assessment, and diagnosis of depression during adolescence. *Psychological Bulletin*, 114, 323-344.
- Diez Zamorano, M. A. (2003). Análisis bibliométrico sobre la depresión infantil en España. *Internacional Journal of Clinical and Health Psychology*, 3, 645-653.
- Donnelly, M. (1995). Depression among adolescents in Northern Ireland. *Adolescence*, 30, 339-350.
- Fristad, M. A., Emery, B. L., and Beck, S. J. (1997). Use and abuse of the Children's Depression Inventory (CDI). *Annals of Clinical Psychiatry*, 3, 341- 346.
- Fristad, M.A., Weller, R.A., Weller, E.B., Teare, M., and Preskorn, S.H. (1991). Comparison of the parent and child versions of the Children Depression Inventory (CDI). *Annals of Clinical Psychiatry*, 3, 341-346.
- Garvin, V., Leber, D., and Kalter, N. (1991). Children of Divorce: Predictors of change following preventive intervention. *American Journal of Orthopsychiatry*, 61, 438-447.
- Gotlib, I.H. (1984). Depression and general psychopathology in university students. *Journal of Abnormal Psychology*, 93, 19-30.
- Gotlib, I.H., Lewinsohn, P.M., and Seeley, J.R. (1995). Symptoms versus diagnosis of depression:

- Differences in psychosocial functioning. *Journal of Consulting and Clinical Psychology*, 63, 90-100.
- Hammen, C. and Rudolph, K.D. (1996). Childhood Depression. In E.J. Mash and R.A. Barkley (Eds.), *Child Psychopathology* (pp. 153-195). New York: Guildford Press.
- Helsen, W.J. and Matson, J.L., (1984). The assessment of depression in children: The internal structure of the Child Depression Inventory (CDI). *Behavior Research and Therapy*, 22, 289-298.
- Joiner, Jr. T., Schmidt, K.L., and Schmidt, N.B. (1996). Low-end specificity of childhood measures of emotional distress: Differential effects for depression and anxiety. *Journal of Personality Assessment*, 67, 258-271.
- Kashani, J.H., Sherman, D. D., Parker, D.R., and Reid, J.C. (1990). Utility of the Beck Depression Inventory with clinic-referred adolescents. *Journal of American Academy of Child Adolescent Psychiatry*, 29, 278-282.
- Kazdin, A. E., Colbus, D., and Rodgers, A. (1986). Assessment of depression and diagnosis of depressive disorder among psychiatrically disturbed children. *Journal of Abnormal Child Psychology*, 14, 499-515.
- Kovacs, M. (1983). *Children's Depression Inventory (CDI)*. Unpublished Manuscript. University of Pittsburg.
- Kovacs, M. (1992). *Children's Depression Inventory (CDI)*. New York: Multi-health Systems, Inc.
- Lobovits, D. A. and Handal, P.J. (1985). Childhood depression: Prevalence using DSM-III criteria and validity of parent and child depression scales. *Journal of Pediatric Psychology*, 10, 45-54.
- Marton, P., Churchard, M., Kutcher, S., and Koremblum, M. (1991). Diagnostic utility of the Beck Depression Inventory with adolescents psychiatric outpatients and inpatients. *Canadian Journal of Psychiatry*, 36, 428-431.
- Mausner, J.S. and Kramer, S. (1985). *Mausner & Bhan, Epidemiology: An introductory text*. West Washington Square Philadelphia: B. Saunders Company.
- Montero, I. and León, O.G. (2005). Sistema de clasificación del método en los informes de investigación en Psicología. *International Journal of Clinical and Health Psychology*, 5, 115-127.
- Nelson, W. M., Politano, P.M., Finch, A. J., Wendel, N., and Mayhall, C. (1987). Children's Depression Inventory: Normative data and utility with emotionally disturbed children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 43-48.
- Nurcombe, B., Seifer, R., Scioli, A., Tramontana, M., Lexington, W., and Beauchesne, H. (1989). Is major depressive disorder in Adolescence a distinct diagnostic entity? *Journal of American Academy of Child-Adolescents Psychiatry*, 28, 333-342.
- Okasaki, S. and Sue, S. (1995). Methodological issues in assessment research with ethnic Minorities. *Psychological Assessment*, 7, 367-375.
- Piotrowski, C., Sherry, D., and Keller, J. W. (1985). Psychoanalytic test usage: A survey of the Society for Personality Assessment. *Journal of Personality Assessment*, 49, 115-119.
- Ramos-Álvarez, M.M. and Catena, A. (2004). Normas para la elaboración y revisión de artículos originales experimentales en Ciencias del Comportamiento. *International Journal of Clinical and Health Psychology*, 4, 173-189.
- Roberts, R. E., Lewinsohn, P. M., and Seeley, J. R. (1991). Screening for adolescents depression: A comparison of depression scales. *Journal of American Academy of Child and-Adolescents Psychiatry*, 30, 58-66.
- Rogler, L. H. (2000). Implementing cultural sensitivity in mental health research: Convergence and New Directions, Part III. *Psychline*, 3, 5-13.

- Rosselló, J., Guisasola, E., Ralat, S., Martínez, S., and Nieves (1992). La evaluación de la depresión en un grupo de jóvenes puertorriqueños. *Revista Puertorriqueña de Psicología*, 8, 155-162.
- Saylor, C. F., Finch, A. J., Spirito, A., and Bennett, B. (1984). The Children's Depression Inventory: A systematic evaluation of psychometric properties. *Journal of Consulting and Clinical Psychology*, 52, 955-967.
- Strober, M., Green, J., and Carlson, G. (1981). Utility of the Beck Depression Inventory with psychiatrically hospitalized adolescents. *Journal of Consulting and Clinical Psychology*, 49, 482-483.
- Tanaka-Matsumi, J. and Kameoka, U.A. (1986). Reliabilities and concurrent validities of popular self-report measures of depression, anxiety and social desirability. *Journal of Consulting and Clinical Psychology*, 54, 328-333.
- Weiss, B. Weisz, J. R., Politano, M., Carey, M., Nelson, W. M., and Finch, A. J. (1991). Developmental differences in the factor structure of the Children's Depression Inventory. *Psychological Assessment*, 3, 38-45.